Computer Science

Assessment of the ACSE Science Learning Environment and the Impact of Movies and Simulations

John F. Pane

1 June 1994
CMU-CS-94-162

Carnegie Mellon

Assessment of the ACSE Science Learning Environment and the Impact of Movies and Simulations

John F. Pane

1 June 1994
CMU-CS-94-162

School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213-3890

Also appears as:
Human-Computer Interaction Institute Technical Report CMU-HCII-94-105

Abstract

This paper describes an empirical study that assesses a multimedia science learning environment, and the impact on student learning of movies and simulations as lesson components. The study measures summative effects on student performance and satisfaction, and gathers formative data about student use of the environment for iterative improvement to the system and lessons. Two lessons containing movies and simulations were compared with lessons that used static graphics to present the same material. Preliminary results show that participants using the lesson with movies and simulations spend significantly more time working through the material, and score higher on questions that target material that is presented with simulation. Analysis of usage patterns identifies features of the system that are effective, underutilized, misunderstood or problematic.

The author can be reached by electronic mail at: pane+acse@cs.cmu.edu

This research is supported by National Science Foundation Grant Number MDR-9150211. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of NSF or the U.S. Government.

Keywords: animation, simulation, multimedia, computer-based learning, science education, human-computer interaction, user studies.

1. Introduction

While traditional teaching methods and materials are very effective in science education, certain skills that are difficult to learn may be facilitated by computer-based learning environments. Two examples are: the understanding of complex time-varying three-dimensional processes; and skills of scientific reasoning, as contrasted with skills of factual recall. The Advanced Computing for Science Education (ACSE) project has designed and built a multimedia science learning environment to address the teaching of these skills [10]. This software system provides the lesson author with a structured document framework and a set of tools that facilitate the construction of science lessons containing text, still graphics, movies, and simulations. The system provides the student with tools for navigating through the lesson, viewing the movies, and manipulating and running the simulations.

We distinguish movies and simulations in the following way: animations, videos, time-lapse photography and other motion pictures that are the same each time they are viewed are classified as movies; and any graphics, tables, or other output generated by running a computer program under student control are classified as simulations. In the latter, the visual results are subject to any changes that the student makes to the simulation program, and thus may be different each time they are viewed.

The educational objective of ACSE is to improve student learning over more traditional teaching materials such as textbooks and lectures. Improvements are hypothesized because certain aspects of ACSE address two perceived needs in traditional science education: 1) movies and simulations, due to their dynamic nature, should improve the understanding of complex time-varying three-dimensional processes; and 2) the investigative nature of the simulations should exercise and improve skills of scientific reasoning as contrasted with skills of factual recall.

Other research has investigated the effect of movies on learning, with mixed results. Studies investigating the use of movies to teach mathematics and dental hygiene have found positive effects [2], [5]. But similar studies in physics learning did not find significant improvements [11], [12]. There is evidence that use of movies may result in speed improvements on immediate recall tasks, but that the effect is reversed on long term recall tasks [9]. But even in cases where improvements are slight, students have a high perceived value of movies [1], [14].

There is evidence that simulation is an effective teaching tool. One study investigated the use of a simulation-based computer-assisted learning system as a supplement to microbiology laboratories. The experimental group achieved a significantly higher posttest score than a laboratory-only control group [6]. Another study compared a traditional lecture class with one that required students to work through a macroeconomic policy simulation. The simulation improved students' ability to solve implicitly stated problems and left them with more positive attitudes toward the discipline [4].

In typical simulation-based learning systems, the program itself is hidden from the user. This is a natural choice, because the science that is imbedded in that simulation is a very small portion of the overall program, and it is not organized in a manner that would be readily understood by the science student [3]. A set of controls are provided that permit the student to manipulate certain well chosen parameters and see the results.

ACSE is distinguished from these typical systems in the way the simulation is integrated with other lesson contents. Selected pieces of the simulation are interspersed throughout the lesson in a manner that is not unlike the use of mathematical formulas in a textbook. These small pieces of the simulation are chosen because they describe the essential science that is embedded within the much larger program. The program elements are provided in the context of explanatory materials, and irrelevant details are hidden.

The student manipulates the simulation by directly modifying these program elements. A full featured novice programming environment is provided to support this activity [7], [13], [8]. Thus the student is afforded the full expressive power of a programming language. This permits the student to investigate the simulation by making structural changes – those that involve modifying or replacing algorithms – in addition to the kinds of parametric changes that are permitted in traditional simulation systems.

This paper describes an empirical study where the full featured ACSE environment is compared with a control environment that is the same except that the movies and simulations are replaced by static graphics that present the same material. These conditions are referred to as *ACSE* and *control*, respectively. The study measures student performance and satisfaction, and gathers formative data to guide future iterations of the system.

2. Description of ACSE Lessons

A lesson in the ACSE environment is called a Volume. The Volume is displayed in a typical Macintosh window with scroll bars along the right and bottom edges (see Figure 1). In addition to a main content area, there are panels at the left edge for a table of contents, a glossary, and a page indicator. A toolbar appears along the top of the window, with tools for navigation. The Volume is divided into pages that fill the content area when the window is sized for a 13 inch monitor.

The study makes use of two different Volumes that were developed to teach topics in developmental biology. The topics, and the materials used to present them, were chosen by a biology professor, whose intended audience was the students in his junior/senior level developmental biology course. The Volumes each take roughly two hours to complete.

The first Volume, entitled *Sea Urchin Gastrulation*, examines a process in early development that is common to many organisms, where cells adopt certain roles and migrate to appropriate locations in the embryo in order to fulfill those roles. This Volume is 51 pages long, and in addition to text contains 15 high resolution images and figures, 4 movies, 6 simulations, 18 fragments of simulation code, and 17 review questions.

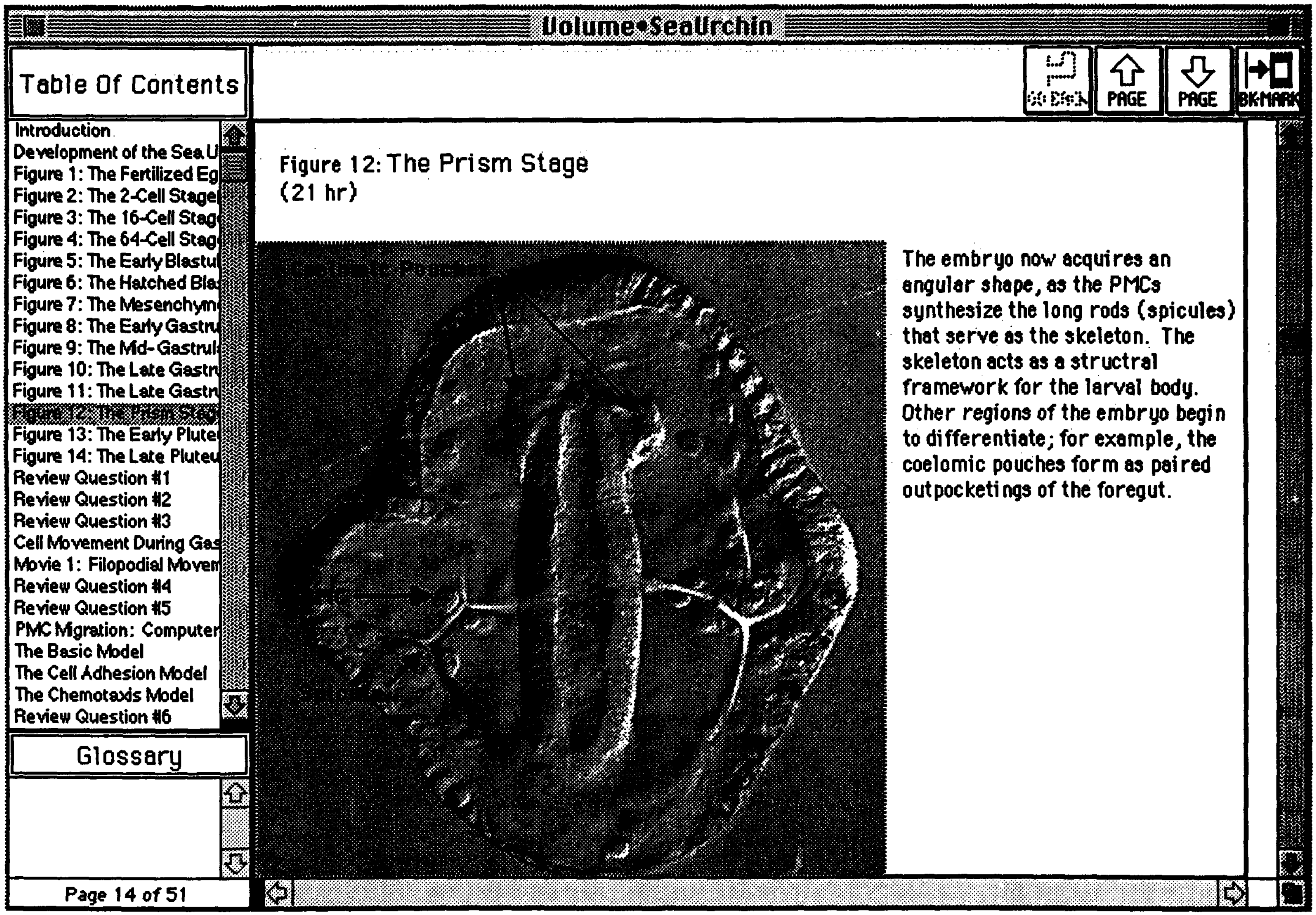

Figure 1: The Volume Window

The second Volume is entitled *Gradients, Gene Expression, and Pattern Formation: The Early Development of Drosophila Melanogaster.* This lesson examines the way that gradients of molecules produced by the mother and stored within the unfertilized egg can result in differences among embryonic cells and the generation of patterns in the organism. This Volume is 55 pages long, and in addition to text contains 23 high resolution images and figures, 3 movies, 7 simulations, 10 fragments of simulation code, and 10 review questions.

The review questions, which are interspersed throughout the Volume, are intended to focus attention on important details, and to encourage active participation in the lesson. The questions are targeted to materials that are presented in each of the following representations: still graphics and text; movies; and simulations. The student is directed to answer the review questions in essay form by typing into the space provided. We will use the answers to these questions as a measure of performance in this study.

3. Experimental Method

3.1. Participants

The participants in the study consisted of 34 students (13 male, 21 female) taking a junior/senior level developmental biology class at Carnegie Mellon University. They participated by attending two scheduled computer laboratory sessions. Their performance during these lab sessions, as measured by their answers to the review questions, constituted 10% of their class grade.

Two pre-trial measures of the participants were gathered: QPA, and scores for the first two exams in the course. Since prior performance in the class was judged to be the best predictor of performance on the volumes, participants were ranked by average exam score, with QPA as a secondary measure for breaking ties. Participants were then assigned to pairs by selecting adjacent students from this ranked list. Within each pair, one student was randomly selected and assigned to Group A; the other was assigned to Group B. Table 1 shows the result of this group assignment. This assignment also resulted in the groups being balanced by sex.

Table 1: Pre-trial Measures: Mean (S.D.)

Group	Exam Scores	QPA
A	68.0 (11.2)	3.08 (0.41)
B	68.0 (11.2)	3.14 (0.43)

Except for one chemistry major (Group B), all of the participants were biology majors. All except one participant (Group A) had completed, or placed out of, at least one university-level computer programming course.

3.2. Materials

All participants were provided with identically configured Macintosh IIsi computers, with 9MB of memory and 13-inch, 8-bit color monitors. When they arrived, the ACSE software was running, and the lesson was pre-loaded. Participants were given a one-page instruction sheet describing the basic features of the ACSE software. Participants answered the review questions by typing directly into the space provided within the lessons.

After finishing work on the lesson, participants were given a short paper questionnaire to answer. The first questionnaire requested some background information with the questions listed in Table 2. These questions were not repeated on the questionnaire that was administered after the second lesson. Except where it was not applicable, the rest of the questions were answered after each lesson. Those questions gathered information about student satisfaction, and comments about the system. The satisfaction questions, listed in Table 3, requested a response on a 5-point scale. Those that requested written comments are listed in Table 4.

Table 2: Survey questions - background information

Have you ever seen anything like this system before?
Do you have a computer at home?
How many hours per week do you use a computer?
What kind of programming experience do you have? (1=none, 5=a lot)

Table 3: Survey questions - participant ratings (1=not at all, 5=very much)

Did you enjoy this session?
Did you learn anything?
Would you ever want to do another lesson again?
Do you think that this system would be good for teaching biology?
Do you like biology?
Do you like computers?
Compared to the last session on the Sea Urchin, how did you like this session? (1=Sea Urchin was better, 5=this was better) [asked after second lesson only]

Table 4: Survey questions – participant comments

What was the worst thing about this experience?
What was the best thing about this experience?
General comments or suggestions as to how we can improve our system:

3.3. Design

The study employed a posttest-only control group design. One factor, the presentation of lesson content, was varied between subjects: the experimental lesson contained movies and simulations; the control lesson did not. Another factor, question type, varied within subjects: three classes of questions – movie-based, simulation-based, and other – were answered by all subjects. The study was conducted in two sessions using two different lessons. The experimental and control groups were reversed for the second session.

3.4. Derivation of Control Lessons

The design of the control condition is modeled after a textbook, but is computer-based in order to minimize confounding factors. Both conditions use the ACSE software, so the only difference between groups is the form of the lesson contents. All elements of the ACSE condition that do not involve movies or simulations appear unchanged in the control condition. Each movie and simulation element is replaced in the control with still graphics that convey the same information, through the process described below.

In order to determine the replacement elements for the first ACSE Volume, we interviewed the lesson designer. For each movie and simulation element, we asked him to reflect on the pedagogical goal of the element by asking questions like: What is the student expected to learn from this Volume element? What are the important events in this time-varying process? Which still images best represent these important events? What captions are required? How would a textbook best describe this material?

The designer's responses to these questions were recorded on videotape, and this tape was reviewed extensively during the derivation process. We then generalized the following criteria in order to guide future derivations, including our derivation of the control for the second ACSE Volume.

3.4.1. Locality and Economy of Space

Biology textbooks typically have strong space limitations, and seldom devote more than two still pictures to a particular topic. Furthermore, it is desirable to maintain locality of information and the approximate total number of pages in the lesson across conditions. Therefore we limit the number of still pictures that are used to replace each movie and simulation to a maximum of four, occupying a maximum of two lesson pages.

3.4.2. Preserve Visual Quality

The visual quality, or information content, of a digital image is dependent on resolution. Two aspects of resolution are spatial resolution and pixel resolution. Spatial resolution quantifies the dimensions of the image, usually expressed in pixels. In addition, each pixel has a resolution, or depth, which determines how many different color and brightness values it may adopt. Pixel resolution is usually expressed in bits/pixel.

One strategy for satisfying the previous principle of locality and economy of space might be to reduce the sizes of the derived still images. This would allow a larger number of images to fit into the allocated space, at the expense of resolution. But, such a manipulation would reduce the information content of the individual images, possibly reducing their effectiveness in showing the desired scientific details. In preparing still images to replace movies and simulations, we prohibit any transformations to the images that affect visual quality through a reduction in spatial or pixel resolution.

3.4.3. Select Key Frames

The limited number of stills that may be chosen to represent a movie or simulation leads to a desire to choose a set that contains the largest amount of relevant information. In most situations, this set is not fixed, but depends on the pedagogical goal of the lesson element. This leads to the following strategy: choose frames representing the initial and final conditions, and intermediate frames that represent important incidents. However, in the event that there is a series of simulations with identical initial conditions, it is not necessary to repeat the frame representing the initial condition each time.

3.4.4. Choose Illustrative Simulation Runs

When students encounter a simulation, they are first directed to run it without modification. Then they are instructed to make a specific modification to the simulation and to run it again, observing the resulting change in output. After these directed runs, they are encouraged to make further modifications to explore the behavior of the system.

Review questions that are targeted to a simulation element are not answered explicitly by the specified simulation runs. Instead, students must make inferences from their observations. In the case of the ACSE group, the inference can be tested by manipulating the simulation and running it. In deriving the control, each of the specified simulation runs is shown. It is essential that these actually illustrate the necessary dynamics, in order that the review question is answerable in the control condition.

3.4.5. Explanatory Text

In deriving the control elements for movies and simulations, it is often necessary to supplement the selected still images with explanatory text. This text might explicitly describe some behavior that is observable in the movie or simulation, or fill in the gaps between the chosen stills. Because this text focuses attention on some pedagogically important aspect of the lesson element which might not otherwise be attended to by the viewer of the movie or simulation, it is also provided in the ACSE lesson.

3.4.6. Replace Program Code with Text

ACSE shows small pieces of the simulation program to the student. Those pieces are selected because they represent a scientifically important aspect of the much larger full simulation. The simulation is written in the Pascal programming language. In some instances, ACSE displays this Pascal code in a natural language form, but in other cases the syntax of the programming language shows through. This programming language syntax may be more or less readable depending on its complexity and the student's programming experience. In the control lesson, simulation code is always translated into natural language or simple mathematical formulas if it is not already displayed that way.

3.5. Procedure

The laboratory sessions were held in a large computer laboratory that is partitioned into two sections. All of the ACSE participants were assigned to one half of the lab; and all of the control participants were assigned to the other. Participants entered this laboratory and checked in with a proctor. The proctor handed the instruction sheet to the participants, and directed them to a computer which had the appropriate lesson opened to page one. The proctor recorded the starting time.

Students had no *a priori* knowledge that there were two different lesson formats, although it may have become apparent in discussions with their peers after the first session. Students were permitted to work as long as they wished. In order to motivate them to take the lesson seriously, they were informed that their answers to the review questions would be graded, and that each lesson was worth 5% of their course grade. There was the additional motivation that the material might appear on an exam. The Volumes were not made available to the students for study outside the laboratory session.

While they were using the system, additional data (e.g. feature usage, timing and keystroke information, etc.) was automatically and unobtrusively gathered by the software. When participants indicated that they were finished, their finishing time was recorded by the proctor, and they were given the questionnaire to fill out.

Due to scheduling conflicts with the scheduled lab sessions, there were six students who attended a makeup session for lesson 1 (four ACSE; two control), and two students who attended a makeup session for lesson 2 (one ACSE; one control). Three students did not participate in lesson 2. The procedure used during these makeup sessions was identical to the main sessions. The only known confounding factor is the possibility that these participants discussed the lesson with their peers before attending the makeup session.

The participants' lesson files were collected by the proctor. Each review question and the participant's answer to that question occupy a full Volume page. So, except for the text of the answer, the contents of these pages do not vary between experimental conditions. These answer pages were printed, and each printout was marked with a unique identifying number that has no relationship to the experimental condition. This procedure ensured that there was no systematic way for an independent grader to determine the author or group assignment of any answer set. However, it is possible that in some cases the contents of a participant's answer revealed this information. Because the grader was not familiar with the details of this study, we do not believe this would bias the grading. When all of the grading is finished, we will debrief the grader to assess this assumption. The grader is a biology graduate student.

4. Results

At the time of this writing, both laboratory sessions have been held, but the review questions for the second lesson have not yet been graded. This section examines preliminary results based on the data available at this time.

4.1. Time on Tas

Table 5 shows the mean time that each group spent working on the lesson. In both cases, the ACSE group spent significantly more time than the control group. On lesson 1, the mean difference between groups was 46 minutes [$t(32) = 4.4$, $p < .001$]. On lesson 2, the mean difference between groups was 26 minutes [$t(29) = 2.1$, $p < .05$].

Table 5: Total Time in Minutes: Mean (S.D.)

Group	Lesson 1	Lesson 2
ACSE	146 (36)	126 (35)
Control	101 (23)	100 (35)

4.2. Location of Time Difference

The first twenty pages of lesson 1 are identical between groups, because there are no movies or simulations on these pages. We measured how long it took each participant to reach page 21, as a rough measure of how long they spent on the first 20 pages. We have not yet checked whether they later returned to these pages and spent more time, or whether they skipped to later pages before reaching page 21. There are no significant differences between groups on this time measure, or on performance on the three questions that appear on those pages. Table 6 shows the group means of each student's average score per question on those three questions.

Table 6: Time and Performance on First Twenty Pages of Lesson 1: Mean (S.D.)

Lesson 1	Time Spent (Minutes)	Score(Percent)
ACSE	34 (16)	73 (14)
Control	27 (11)	76 (13)

4.3. Questionnaire Results

Table 7 shows the responses to the participant rating questions that were listed above in Table 3. Except where noted, all of these were answered on a 5-point scale (1=not at all, 5=very much). There are no significant correlations between these participant ratings and group or time spent.

Table 8 shows the responses to the background questions that were listed above in Table 2. There are no significant differences between groups on these responses.

Table 7: Participant Ratings: Mean (S.D.)

Question	Lesson 1		Lesson 2	
	Group A	Group B	Group B	Group A
	ACSE	Control	ACSE	Control
Enjoy this session?	3.9 (0.9)	3.8 (1.0)	2.8 (0.9)	2.8 (1.3)
Learn anything?	3.6 (1.0)	4.0 (0.9	3.7 (1.1)	3.1 (0.9)
Want to another lesson?	3.5 (1.1)	3.8 (0.8)	3.0 (1.1)	2.8 (1.4)
Good for teaching biology?	4.1 (0.9)	4.5 (0.7)	3.7 (1.0)	3.3 (1.2)
Like biology?	4.8 (0.6)	4.7 (0.5)	4.4 (0.9)	4.5 (1.1)
Like computers?	4.1 (1.0)	3.5 (0.9)	3.4 (1.1)	4.1 (0.8)
Compared to previous lesson, how much did you like this? (1=lesson 1 better, 5=lesson 2 better)			2.1 (1.0)	2.7 (1.9)

4.4. Participant Comments

The most common complaint on the questionnaires was that the lesson took too long or was too slow. This was mentioned 40 times on ACSE surveys, and 7 times on control surveys. In addition, on lesson 2 there were three students who commented that the best thing about les-

Table 8: Participant Background Information

Question	Group A	Group B
Seen anything like this [% yes]	41%	13%
Computer at home [% yes]	53%	65%
Computer hours per week [mean (S.D.)]	13.1 (7.6)	12.1 (8.5)
Programming Experience (1=none,5=a lot) [mean (S.D.)]	2.9 (1.0)	2.6 (0.9)

son 2 is that it is faster than lesson 1. All of these students were in the ACSE group on lesson 1 and the control group on lesson 2.

The most common positive comments regarded the quality of the visual aids. This was mentioned 22 times by the ACSE group, and 20 times by the control group. Of these, 10 of the ACSE comments specifically mentioned movies and simulations. There were five comments (1 ACSE, 4 control) that viewing the 3d images strained the eyes. [1]

On lesson 1, there were two comments from the control group about the lack of interactivity or suggesting that the lesson include movies and simulations. On lesson 2, this comment was received nine times from control subjects. Recall that these are the students who had seen and used movies and simulations on lesson 1.

One control participant commented that it was difficult to answer "the thought-provoking questions." One control participant said that the simulation-based questions were better because of the thought process involved.

4.5. Performance on Review Questions

In addition to measuring the overall score on the lesson 1 review questions, we calculated partial scores for three categories of questions: movie-based; simulation-based; and other. For each question, this classification was decided by locating the Volume element that is most useful for answering the question, and determining whether that element is (in the ACSE condition) a simulation, a movie, or some other type of element. The results are shown in Table 9. The numbers shown represent the average score on questions in the categories listed, expressed as a percentage of the possible points in that category. A significant difference between groups was found only on simulation-based questions. On those questions, the ACSE group scored about 10% higher than the control group [$t(32) = 2.153$, $p < .05$].

1. The 3d images are viewed with cardboard glasses containing one red and one blue lens.

Table 9: Performance on Review Questions by Question Type on Lesson 1: Mean (S.D.)

Group	Overall (17)	Movie (5)	Simulation (5)	Other (7)
ACSE	73 (9)	70 (12)	78 (10)	72 (10)
Control	71 (9)	72 (14)	71 (9)	71 (10)

4.6. Pretest Measures as Performance Predictors

We found no significant correlation between the pretest measures (QPA, Prior Exams) and performance overall or performance on any of the categories of questions.

5. Discussion

These preliminary results guide the following discussion of the location of the time differences, the kinds of skills that were enhanced, and the magnitude of those improvements.

5.1. Time Differences

In both lessons, the control Volume is approximately the same length as the ACSE Volume. However, the process used to derive the control ensures that it has no more total information than the ACSE lesson. This, in combination with the fact that each ACSE simulation takes several minutes to run and can be run an unlimited number of times, suggests that the ACSE group will take more time to work through the material. This is confirmed by our observations of total time.

Analysis of the first 20 pages of lesson 1, which are the same between groups, supports the expectation that any time or performance difference between groups is due to the experimental manipulation that appears later in the lesson. This analysis is approximate because we simply measured the elapsed time until the participants reached page 21. Complete analysis will measure whether any of that time was spent skipping ahead without hitting page 21, or any additional time that was spent in subsequently returning to the first 20 pages.

Participants might have spent more time on a lesson if they were enjoying the experience or felt that it was a productive use of their time. Alternatively, participants might have felt some obligation to complete the lesson, and thus would have been less satisfied if it was not enjoyable or seemed an unproductive use of time. Since the student satisfaction measures do not show significant correlations to time spent, neither of these speculations can be confirmed.

5.2. Skills of Scientific Investigation

While recall of facts is an important part of science education, skills of scientific investigation are perhaps even more valuable. To some extent, the simulation-based questions target these skills. Many of them take one of two general forms: 1) If the biological system behaved in manner X, what would happen? and 2) What must be true about the biological system in order for Y to happen? These questions lead the students on an investigation that resembles the research that discovered the workings of these biological systems. A series of appropriate and meaningful questions are posed to isolate important aspects of the biological system.

An even more valuable skill would be for the student to take the initiative of the investigation. There is one question on lesson 2 that attempts to measure this skill by asking the student *both* to pose a meaningful question about the system and to describe a scientific method for answering it. If the ACSE environment improves students' skills of scientific investigation, we would expect to see an even greater effect between groups on this question.

5.3. Magnitude of Performance Improvement

There was no significant overall performance improvement in the ACSE group. The approximately 10% improvement on the simulation questions, which was statistically significant, appears to be a small effect on which to judge the success of the ACSE system design. However, different methods of testing the system may yield larger effects.

The magnitude of the performance improvement should be considered in the context of the small amount of time allotted to each lesson. In addition, students received no instruction on the features and effective use of the system. Thus this study does not test the effectiveness of this system if it were to be adopted as the central focus of a course, where students and instructor use the system daily.

It may be the case that even the control lesson in this study was superior to the other teaching materials that are available to the instructor. It is derived from a carefully constructed lesson that covers exactly the topics desired, in exactly the order that this particular instructor would choose to present them. It uses still graphics that were generated by a simulation which would otherwise not be available were it not for the ACSE environment. This is analogous to writing a custom textbook for the course. As a rough comparison of this class with one that was taught by this instructor before the introduction of ACSE Volumes into the course, we have requested that he review old exams for relevant questions that can be reused with the current class. To the extent that he was dissatisfied with student performance on these past questions, we can test whether the current students show any improvement.

6. Formative Evaluation

In addition to direct feedback from the participants on the questionnaire, the software logs contain records of usage patterns and problems. These following observations about the system were facilitated by these logs.

6.1. Navigation

The design of the ACSE system provides a table of contents along the left edge of the page. This feature shows progress through the lesson, and allows quick navigation from one section to another. This feature was not implemented for lesson 1. We received four questionnaire comments (1 ACSE, 3 control) requesting this feature or something like it. The feature was implemented for lesson 2, and the software logs show that it was heavily used. Three participants (1 ACSE, 2 control) reported that the table of contents was the best thing about their experience with lesson 2.

However, not every page of the lesson has a table of contents entry, so some other method of scrolling must be used to reach those pages. One student exclusively used the table of contents for navigation, thus skipping pages of the lesson that do not appear there. This happened even though this student had used the other scrolling mechanisms on lesson 1, there is a page number indicator visible at all times, and the text had several cross references to these pages. We will investigate whether this participant's problem is an anomaly, or a genuinely confusing aspect of the current software design. We will explore remedies, such as annotating each table of contents entry with a page number.

6.2. Management of Information

There were six comments (1 ACSE, 5 Control) about difficulty in remembering material from page to page. Page traces during times when students are answering the review questions show that they often scroll back through the material multiple times in the midst of typing their answers. This problem may be relieved when planned features – a glossary and live cross-reference links – are completed. One subject suggested that the system provide a scratchpad that the user can use to record notes and simulation results. We will consider implementing this, perhaps in the form of an online lab notebook.

6.3. Programming

In the ACSE software, we experiment with an alternate syntax for presenting simulation code. One example is the display of procedure calls, which in the Pascal syntax look like:

```
Regulator(bicoid,1000,activates,hunchback,50);
```

These are instead displayed in sentence form:

Regulator: when the concentration of **bicoid** is greater than **1000** units, it **activates** the expression of **hunchback** with an efficiency of **50%**.

This special formatting poses some problems when editing is initiated, because the underlying Pascal syntax appears at that time, and the user must enter changes using that syntax. The natural language syntax reappears when editing is completed. The simulation logs show that this transition between representations was a problem for many students: they encountered syntax errors and sometimes inadvertently destroyed the annotations that provide the natural language syntax. So, while this natural language syntax may be more readable, its benefits may be cancelled by the disruptive changes in syntax that occur on entry and exit from editing. We will investigate ways to remedy this problem while retaining the readability of the natural language syntax.

We also observed students having difficulty with the Pascal syntax in other cases: 1) real numbers must have a digit before the decimal place, so values like ".8" are not accepted; and 2) commas are not permitted in large numbers, so "32,000" must be entered as "32000". These problems highlight the limitations of using a programming language as a language of scientific discourse. Buttons and sliders, which are a planned but not yet implemented feature of the software design, can be used to provide a direct-manipulation interface to numeric parameters. This class of problem can also be mitigated by deviating from the Pascal language standard to selectively permit incorrect but unambiguous syntax to be interpreted in the expected way.

6.4. Workstation Requirements

The most frequent positive and negative comments have implications about the workstation requirements of this kind of learning system.

The predominant positive comment from both groups concerned image quality. The ACSE lessons make extensive use of large, high resolution color graphics. It would be inappropriate to attempt to use these lessons on systems that do not support color graphics or that have miniature screens. The memory and processing demands of these high-quality graphics were at the upper limit of the capabilities of the systems that were used in this study.

The most common negative comment regarded speed, and this was biased heavily toward users of the ACSE lesson. This can be explained by the processing requirements of the simulations, which are present only in the ACSE condition. In many cases the simulations took several minutes to run on the participants' computers. This indicates that the system was underpowered for this task. Clearly it would be unwise to attempt to use these lessons on less powerful machines. There are now machines available that are orders of magnitude faster than, but in the same price range as, the four-year-old systems that were used in this study.

7. Future Work

These preliminary results will direct further analysis of the data collected in this evaluation, and will guide further development of the ACSE software and further investigation of the effective use of the system.

7.1. Source of Improvements

The measured improvement on simulation-based questions is encouraging, and merits further investigation. In addition to analyzing the scores on the second lesson when they become available, we will use the software logs to investigate the ways that simulations were used by participants, and how that impacted their performance on the corresponding review questions. We will determine whether only the suggested simulations were run, or whether participants generated their own situations and tested them. We will look at the amount of time they spent with the simulations, and the number of times they ran them.

The improved performance on the simulation questions can be partially accounted for by the extra total time spent by the ACSE group. We will study the correlation on individual questions between score and time spent.

7.2. Investigating Learning During Long-Term Usage

The design of the current study does not provide a measure of long-term learning. We intend to investigate this, due to the previously mentioned study showing that a positive effect of movies may disappear or become a negative effect in the long term [9]. This environment can test whether this also occurs in simulation-based learning. In the near term, we will gather results from any questions from the final exam that are relevant to either of these lessons, and see whether there is any measurable longer-term effect of ACSE.

Early analysis of the software logs shows that the participants did not make major structural changes to the simulations. This is not surprising, given the short amount of time that they actually spent with the system. In future studies we would like to test whether students make more effective use of the simulations when they have a longer period of time, such as a week, to work with the lesson. In addition, careful design of the lesson may increase the likelihood that these structural changes will be attempted by the student.

The instructor of the course is very enthusiastic about teaching with the ACSE environment. It addresses frustrations he has had in teaching about complex time-varying three dimensional processes. If we had lessons that covered his entire developmental biology course, he would make use of them. We could then ask whether students using the ACSE environment many times throughout the semester learn to use it more effectively.

8. Conclusions

The preliminary results of this study suggest an encouraging conclusion about the impact of simulation on student learning. While students spend significantly more time with lessons that include simulation and movies, they score better on questions that target scientific investigation skills than students in a control group that lacks these features. We plan to analyze further the source of this improvement and investigate ways that simulation can most effectively be used in teaching science. In addition to these summative effects, the study provides useful formative guidelines for future iterations of the software.

9. Acknowledgments

The project is a collaboration among Carnegie Mellon University's Department of Computer Science, Department of Biological Sciences, and Center for Light Microscope Imaging and Biotechnology, and Human Computer Interaction Institute. The author acknowledges the extensive contributions of Al Corbett, Bonnie John, and Ken Koedinger, who spent countless hours assisting in the design and analysis of this study. The ACSE project team had important input into all aspects of the ACSE system, the lessons, and this study. In particular: Laurie Damianos developed the ACSE and Control Volumes into the excellent lessons that they are today, and organized and proctored the laboratory sessions; Glenn Meter played a major role in the design of the ACSE system, and implemented that design almost single-handedly; Phil Miller inspired and directed the ACSE project, and provided invaluable feedback by adopting the ACSE environment in his classes; Scott Vorthmann developed many of the advanced features of the underlying programming environment, as well as some of the simulations; and Chuck Ettensohn provided the topics and materials for the lessons, and was willing to test them in his class.

10. References

[1] Badre, A., Beranek, M., Morgan Morris, J., Stasko, J. (1992). Assessing program visualization systems as instructional aids. *Proceedings of 4th International Conference on Computer Assisted Learning (ICCAL '92)*, Wolfville, Canada, June 1992, pp 87-99.

[2] Beak, Y.K., & Layne, B.H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction,* 15 (4), pp 131-135.

[3] Fazarinc, Z. (1988). Overhead in writing physics courseware. *Proceedings of the 1988 Asia-Pacific Conference on Computer Education,* Shanghai, China, pp 428-431.

[4] Grimes, P.W., & Willey, T.E. (1990). The effectiveness of microcomputer simulations in the Principles of Economics course. *Computers & Education,* 14 (1), pp 81-86.

[5] Lautar-Lemay, C. (1987). Comparison of computer-assisted-learning with traditional lecture and reading assignment. *Proceedings of the International Conference on Computer Assisted Learning in Post-Secondary Education,* Calgary, Canada, May 1987, pp. 293-294.

[6] Lazarowitz, R., & Huppert, J. (1993). Science process skills of 10th-grade biology students in a computer-assisted learning setting. *Journal of Research on Computing in Education,* 25 (3), pp 366-382.

[7] Miller, P.L., & Chandhok, R.P. (1989). The design and implementation of the Pascal Genie. *Proceedings of the 1989 ACM Computer Science Conference,* Louisville, KY.

[8] Myers, B.A., Chandhok, R.P., and Sareen, A. (1988). Automatic data visualization for novice Pascal programmers. *Proceedings of the 1988 IEEE Workshop on Visual Language,* Pittsburgh, PA, pp 192-198.

[9] Palmiter, S., Elkerton, J., & Baggett, P. (1991). Animated demonstrations vs. written instructions for learning procedural tasks: a preliminary investigation. *International Journal of Man-Machine Studies,* 34 (5), pp 687-701.

[10] Pane, J.F., & Miller, P.L. (1993). The ACSE multimedia science learning environment. *Proceedings of the 1993 International Conference on Computers in Education,* Taipei, Taiwan, pp 168-173.

[11] Rieber, L.P. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research,* 5 (4), pp 431-444.

[12] Rieber, L.P., Boyce, M.J., & Assad, C. (1990). The effects of computer animation on adult learning and retrieval tasks. *Journal of Computer-Based Instruction,* 17 (2), pp 46-52.

[13] Roberts, J., Pane, J., Stehlik, M., & Carrasquel, J. (1988). The Design View, a high level visual programming environment. *Proceedings of the 1988 IEEE Workshop on Visual Language,* Pittsburgh, PA, pp 213-220.

[14] Stasko, J., Badre, A., & Lewis, C. (1993). Do algorithm animations assist learning? An empirical study an analysis. *Proceedings of ACM INTERCHI'93 Conference on Human Factors in Systems,* Amsterdam, Netherlands, April 1993, pp 61-66.